



**HAL**  
open science

## Classification and Regression Trees for Bacterial Vaginosis Diagnosis in Pregnant Women Based on High-Throughput Quantitative PCR

Aurore Loquet, Rémi Le Guern, Teddy Grandjean, Claire Duployez, Marvin Bauduin, Eric Kipnis, Gilles Brabant, Damien Subtil, Rodrigue Dessen

► **To cite this version:**

Aurore Loquet, Rémi Le Guern, Teddy Grandjean, Claire Duployez, Marvin Bauduin, et al.. Classification and Regression Trees for Bacterial Vaginosis Diagnosis in Pregnant Women Based on High-Throughput Quantitative PCR. *Journal of Molecular Diagnostics*, 2021, 23 (2), pp.234-241. 10.1016/j.jmoldx.2020.11.004 . hal-03311394

**HAL Id: hal-03311394**

**<https://hal.science/hal-03311394v1>**

Submitted on 13 Feb 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

## **Classification and regression-trees for bacterial vaginosis diagnosis in pregnant women based on high-throughput quantitative PCR**

Aurore LOQUET<sup>1</sup>, Rémi LE GUERN<sup>1-2</sup>, Teddy GRANDJEAN<sup>1</sup>, Claire DUPLOYEZ<sup>1-2</sup>, Marvin BAUDUIN<sup>1</sup>, Eric KIPNIS<sup>1-3</sup>, Gilles BRABANT<sup>4</sup>, Damien SUBTIL<sup>5-6</sup>, Rodrigue DESSEIN<sup>1-2</sup>.

- 1- Univ. Lille, CNRS, UMR9017 Inserm U1019, Institut Pasteur de Lille, CIIL - Center for Infection and Immunity of Lille, F-59000, Lille, France
- 2- CHU Lille, Institut de Microbiologie, F-59000 Lille, France
- 3- CHU Lille, Service de Réanimation Chirurgicale, F-59000, Lille, France
- 4- Groupement des Hôpitaux de l'Institut Catholique de Lille, Hôpital Saint Vincent, F-59000 Lille, France
- 5- Univ. Lille, CHU Lille, Pôle Femme Mère Nouveau-né, F-59000 Lille, France
- 6- Univ. Lille, EA 2694 : épidémiologie et qualité des soins, F-59000 Lille, France

**Running Title:** Diagnosis of vaginosis in pregnancy

### **Corresponding Author**

Pr. Rodrigue Dessein, Univ. Lille, CNRS, UMR9017, Inserm U1019, CHU Lille, Institut Pasteur de Lille, CIIL - Center for Infection and Immunity of Lille, F-59000, Lille, France.

e-mail : [rodrigue.dessein@chru-lille.fr](mailto:rodrigue.dessein@chru-lille.fr);

**Funding:** Supported by the Ministry of Health and Research (grant PHRC2004/1918 PROM04-06-859 to R.D., D.S., and G.B.).

**Disclosures:** None declared.

## ABSTRACT

Bacterial vaginosis (BV) diagnosis in pregnancy is based on the Nugent score, which consists of semi-quantitation of bacterial morphotypes after Gram coloration. There are limited data concerning molecular method-based diagnosis in asymptomatic pregnant women. Using high-throughput quantitative PCR (qPCR), 34 microorganisms were screened in asymptomatic pregnant women compared to the Nugent score. Three hundred and four vaginal samples had a Nugent score  $<7$  (69.9%) and 131 a Nugent score  $\geq 7$  (30.1%), consistent with BV. More pregnant women with BV share *A. vaginae*, *Bacterial Vaginosis Associated Bacteria-2*, *Gardnerella* spp., *M. curtisii*, *M. mulieris*, *M. hominis*, *U. urealyticum*, *P. bivia*, *Megasphaera 1*, *Megasphaera 2* in their vaginal sample. Fewer pregnant women with BV share *L. crispatus*, *L. gasseri*, *L. jensenii* and *E. faecalis* in their vaginal sample ( $p < 0.001$ ). Then, classification and regression-trees (CART) analysis was performed to determine which combinations of detected bacteria optimally diagnose BV in this population. A set of only four bacteria out of 34 microorganisms (*A. vaginae*, *Gardnerella* spp., *L. crispatus*, and *P. bivia*) was the best combination to identify BV in a cohort of asymptomatic pregnant women with sensitivity to 77.1%, and the specificity to 97.0% compared to Nugent score. The qPCR in the present study responds to the limits of the Nugent score by implementing an easy reproducible quantitative assay to assess the absence of BV in pregnancy.

## INTRODUCTION

Bacterial vaginosis (BV) is characterized by a dramatic shift in the vaginal microbiota population, which involves the loss of beneficial bacteria (normally *Lactobacillus*-dominated) and the simultaneous proliferation of both anaerobes and genital mycoplasmas<sup>1</sup>. During pregnancy, BV increases the risk of preterm birth (odds ratio, 2.19; 95% CI, 1.54-3.12)<sup>2</sup>. This risk increases when BV is present before 20 weeks of gestation (odds ratio, 4.20; 95% CI, 2.11-8.39), and increases even further when BV is present before 16 weeks of gestation (odds ratio, 7.55; 95% CI, 1.80-31.65)<sup>2</sup>. Additionally, BV also increases the risks of spontaneous abortion (odds ratio, 9.91; 95% CI, 1.99-49.34) and of maternal infection (odds ratio, 2.53; 95% CI, 1.26-5.08<sup>2</sup>).

BV may be diagnosed non-microbiologically according to Amsel's criteria: the presence of clue cells, vaginal pH higher than 4.5, profuse white discharge, and fishy odor following discharge exposition to potassium hydroxide<sup>3</sup>. Microbiological diagnosis consists of the Nugent score established by Gram staining of a vaginal smear<sup>4</sup>. The Nugent score consists of semi-quantitation of the following bacterial morphotypes: *Lactobacillus* species (Gram-positive rods), *Gardnerella spp* (Gram-negative rods), and *Mobiluncus* (curved Gram-variable rods). The Nugent score classifies vaginal smears into: healthy microbiota (score 0-3), intermediate or mixed microbiota (score 4-6), and BV (score 7-10). The Nugent score is more specific for BV and more reproducible than Amsel's criteria, especially during pregnancy<sup>5</sup>. The Nugent score has been used as a gold standard in pregnant women for BV diagnosis in many randomized controlled trials evaluating the impact of BV treatment on prematurity<sup>6</sup> and the Nugent score remains considered a gold standard for screening BV in pregnant women by the French High Authority of Health ([https://www.has-sante.fr/upload/docs/application/pdf/prevention\\_antenatale\\_du\\_risque\\_infectieux\\_bacterien\\_-\\_syn.pdf](https://www.has-sante.fr/upload/docs/application/pdf/prevention_antenatale_du_risque_infectieux_bacterien_-_syn.pdf); [French] last update 19/07/2006).

Molecular methods present several theoretical advantages compared to Gram-staining<sup>7,8</sup>. Molecular methods can detect fastidious bacteria or bacteria not seen on microscopic examination. While molecular methods seem to perform well in the diagnosis of symptomatic BV, there is limited data concerning the molecular method-based diagnosis of BV in pregnant women in general, and asymptomatic parturients in particular<sup>8,7,9</sup>. Furthermore, since commercial molecular methods screen a limited number of different bacteria associated with BV, it remains unclear how many or which bacteria are sufficient to diagnose BV in asymptomatic pregnant women.

Using the commercially available high-throughput quantitative PCR Open-Array platform (ThermoFisher Scientific, Waltham, MA, USA), 34 microorganisms were screened in a cohort of pregnant women in comparison with Nugent score screening. Classification and regression-trees (CART) analysis was performed to determine which combinations of detected bacteria optimally diagnose BV in this population.

## MATERIALS AND METHODS

### Study design

The present work is an ancillary study of the PREMEVA2 study (Study of the risk of prematurity according to the balance of bacterial vaginal microbiota; <https://www.clinicaltrials.gov/>, NCT03618836). A total of 456 vaginal samples of pregnant women came from the prospective observational PREMEVA2 study sample collection. The inclusion criteria were gestational age less than 15 weeks, a maternal age of 18 years or older, and the ability to speak French and provide written informed consent. The exclusion criteria were vaginal bleeding within the week prior to proposed screening of bacterial vaginosis. Pregnant women over 18 years old were included after written informed consent. Included women accepted to provide a vaginal self-sample during their first trimester of pregnancy for research purposes. We performed Gram-staining and Nugent scoring on vaginal smears, and the remaining sample was stored at -80°C. The regional institutional review board, which follows the Declaration of Helsinki and the guidelines for Good Clinical Practice, approved the study (Comité de Protection des Personnes participant à la Recherche Biomédicale, PHRC2004/1918 PROM04-06-859).

Pregnant women in the Nord-Pas de Calais region of France (participating hospitals and private gynecology consults) were offered free screening for bacterial vaginosis during their first trimester of pregnancy on self-collected vaginal samples and proposed to participate in banking one sample of vaginal fluid. The vaginal smear was then Gram-stained and 10-20 microscopic fields were examined at 1,000x magnification (with oil immersion) to assess bacterial and cellular abundance. Slides were categorized according to Nugent scores: 0–3 considered normal (*Lactobacillus*-dominant), 4–6 considered intermediate (mixed morphotypes), and 7–10 indicative of bacterial vaginosis (absence of *Lactobacilli* and predominance of the other two morphotypes)<sup>10</sup>.

## **Vaginal microbiota detection by Open-Array high-throughput qPCR**

DNA was extracted from vaginal samples using KingFisher Flex Magnetic Particle Processor and the MagMAX DNA Multi-Sample Ultra Kit (ThermoFisher Scientific). OpenArray is a commercial high-throughput, nanofluidic real-time PCR platform. OpenArray plates were custom designed by Life Technologies Bioinformatics group (ThermoFisher Scientific). A fluorescent dye (ROX), in the TaqMan assay mixture, provided a quality assessment of the manufacturing and loading of the wells. The primer and probe sets targeting bacteria were pre-spotted in triplicate on each subarray. Primer and probe sets are proprietary, designed by the manufacturer, and non-disclosed (ThermoFisher Scientific).

Due to limitations in the capacity of current molecular methods to discriminate between *Gardnerella leopoldii*, *Gardnerella piotii*, *Gardnerella swidsinskii* and *Gardnerella vaginalis*, within the *Gardnerella* spp. genus, *Gardnerella* spp. is referred throughout our results despite the TaqMan assay referring to *G. vaginalis* in amplification results according to manufacturer specifications<sup>11,12</sup>.

An individual TaqMan assay was then loaded in each well-containing 33nL of PCR reaction mixture. Purified DNA from vaginal samples were transferred to OpenArray plates using the automated OpenArray AccuFill system (ThermoFisher Scientific).

OpenArray plates were loaded on the QuantStudio 12K thermocycler (ThermoFisher Scientific). The amplification program consisted of 50°C for 15 seconds, 91°C for 10 minutes, followed by 50 cycles of 54°C for 170 seconds, and 92°C for 45 seconds. The limit of detection was 100 copies per reaction for the 16S rDNA gene. The annealing temperature was 60 °C. Clinical evaluation of the platform showed an overall sensitivity of 93% and a specificity of 90% compared to gold-standard testing<sup>13</sup>.

Quantification standard curves were added with a multi-target plasmid harboring each target sequence. Thirty-four bacteria or fungi or viruses were screened: *Trichomonas vaginalis*, *Chlamydia trachomatis*, *Neisseria gonorrhoeae*, *Herpes Simplex Virus-2* (HSV2), *Herpes Simplex Virus-1* (HSV1), *Treponema pallidum*, *Haemophilus ducreyi*, *Enterococcus faecalis*, *Streptococcus agalactiae*, *Staphylococcus aureus*, *Escherichia coli*, *Gardnerella spp.*, *Atopobium vaginae*, *Bacterial Vaginosis Associated Bacteria-2* (BAVB2), *Mycoplasma hominis*, *Prevotella bivia*, *Bacteroides fragilis*, *Ureaplasma urealyticum*, *Mobiluncus curtisii*, *Mobiluncus mulieris*, *Mycoplasma hominis*, *Mycoplasma genitalium*, *Megasphaera 1*, *Megasphaera 2*, *Lactobacillus iners*, *Lactobacillus crispatus*, *Lactobacillus gasseri*, and *Lactobacillus jensenii*, *Candida albicans*, *Candida glabrata*, *Candida lusitanae*, *Candida krusei*, *Candida parapsilosis*, *Candida tropicalis*, *Candida dubliniensis*.

### **Quality control**

We performed calibration curves for all the tested microorganisms in each plate to guarantee microbial quantification. Additionally, amplification of prokaryotic 16S rRNA and human Ribonuclease P RNA P Component 1 (RPPH1) genes were used as an internal control of the extraction process. Negative control wells ensured that no contamination occurred. QuantStudio 12K Flex software was used to analyze open array data. The following controls were used to verify the quality of amplification: absence of irregularity in the amplification curve, a standard deviation  $\leq 0.5$ , a Cq Confidence  $\geq 0.8$ . Results were expressed as copies of microorganism DNA per mL.

### **Statistical analysis**

Statistical analysis was performed with R software version 3.6.3. Continuous variables were compared with Student's t-test after log transformation. Classification and regression trees (CART) analysis was used for predicting modeling, using the R package caret (Classification And REgression Training) (R software, version 3.6. 1). Cross-validation was included in the



model-building algorithm using k-fold cross-validation ( $k = 10$ ). k-fold cross-validation was repeated ten times to create many versions of the folds and aggregates the results. Performance characteristics of the model were calculated (area under curve, sensitivity, specificity, positive predictive values, negative predictive values).

## RESULTS

Thirty samples were excluded due to the absence of human RNase (n = 11) or 16S rDNA (n=22) amplification. Three hundred and four vaginal samples had a Nugent score <7 (69.9%) and 131 a Nugent score  $\geq$ 7 (30.1%), consistent with BV. The median gestational age of women with a Nugent score <7 was 65 days [IQR 48; 80], versus 61 days [IQR 46;81] for women with BV (p=0.61). The median age of women with a Nugent score <7 was 29 years [IQR 25; 31], versus 27 [IQR 24; 31] for women with BV (p=0.017). Sexually-transmitted bacteria were detected in 15/435 women (3.4%) (*C. trachomatis*, n=10; *T. vaginalis*, n=4; HSV2, n=1) (Table 1). Yeasts belonging to *Candida* spp. were detected in 66/435 women (15.2%); 54 of these were *Candida albicans*. No statistical difference was observed between Nugent $\geq$ 7 or Nugent <7 groups for all the microorganisms tested.

Qualitative analysis showed that *A. vaginae*, BVAB2, *Gardnerella* spp., *M. curtisii*, *M. mulieris*, *M. hominis*, *U. urealyticum*, *P. bivia*, *Megasphaera 1*, *Megasphaera 2*, *E. coli* and *S. agalactiae* were more detected in pregnant women with BV (Nugent score  $\geq$ 7). Conversely, *L. crispatus*, *L. gasseri*, *L. jensenii*, *E. faecalis* were less frequently detected in pregnant women with BV (Nugent score  $\geq$ 7) (Table 1).

Quantitative amplification showed a higher concentration of *A. vaginae*, BVAB2, *Gardnerella* spp., *M. curtisii*, *M. mulieris*, *M. hominis*, *U. urealyticum*, *P. bivia*, *Megasphaera 1*, *Megasphaera 2* and *S. agalactiae* in pregnant women with BV (Nugent score  $\geq$ 7). *E. coli*, *E. faecalis*, *L. crispatus*, *L. gasseri*, *L. jensenii* were lower in pregnant women with BV (Nugent score  $\geq$ 7) (Figure 1). No significant differences were observed for *B. fragilis*, *C. trachomatis*, *L. iners*, *S. aureus*, *C. albicans*, *C. glabrata* and *T. vaginalis* (Figure 2).

Classification and Regression Trees (CART) analysis was used to the optimal bacterial detection paths to BV diagnosis (13). *A. vaginae*, *Gardnerella* spp., *L. crispatus*, and *P. bivia* were the four bacteria that maximize the diagnostic model (Figure 3). CART analysis shows that *A. vaginae* over 6.4 log copies/mL accurately diagnoses BV in 17% of the asymptomatic

pregnant women. *L. crispatus* over 2.4 log copies/mL is associated with an absence of BV in 43% of the asymptomatic pregnant women in the context of low loads of both *Gardnerella spp.* (<6.4 log copies/mL) and *A. vaginae* (<6.4 log copies/mL). *P. bivia* could discriminate between the absence or presence of BV among 31% of the studied population in which *A. vaginae*, *Gardnerella spp.*, and *L. crispatus* are all decreased. The area under the curve for this approach was 0.93. Overall accuracy was evaluated at 91.0%, sensitivity 77.1%, specificity 97.0%, predictive positive value 91.8%, and negative predictive value 90.8%.

## DISCUSSION

A set of only four bacteria out of 34 microorganisms (*A. vaginae*, *Gardnerella* spp., *L. crispatus*, and *P. bivia*) was the best combination to identify BV in a cohort of asymptomatic pregnant women using a commercially available high-throughput quantitative PCR. A small proportion of pregnant women had sexually transmitted microorganisms that did not interfere with the classification of BV. In the studied cohort of asymptomatic pregnant women, a CART algorithm individualizes homogenous sub-groups that are relevant to the pathogenesis of BV.

Previous studies using qPCR have shown that high loads of *A. vaginae* combined with *Gardnerella* spp. allow the diagnosis of symptomatic BV in non-pregnant women<sup>14,15</sup>. Moreover, in a cohort of pregnant women symptomatic or not for BV, Menard *et al.* reported the highest predictive value for BV diagnosis with a molecular quantification of *A. vaginae*  $\geq 10^8$  copies/mL and *Gardnerella* spp.  $\geq 10^9$  copies/mL<sup>16</sup>. However, a total of 51 vaginal samples were positive to both bacteria in the 101 pregnant women with a Nugent score  $\geq 7$  (50%) and a total of 83 vaginal samples were positive to both bacteria among the 304 pregnant women with a Nugent score  $< 7$  (27%). These results explain why CART analysis does not require the combination of *A. vaginae* with *Gardnerella* spp. to achieve BV diagnosis in all pregnant women asymptomatic for BV. The role of *G. vaginalis* is controversial, since this bacterium is also present in 10–40% of healthy women<sup>17</sup>. Regardless, *G. vaginalis* remains a key bacterium that allows diagnosis of BV in the present study. This discrepancy may stem from recent evidence suggesting that *G. vaginalis* biofilms instead of dispersed cells are a marker of BV<sup>18,19</sup>. Some strains of *G. vaginalis* seem to have a significantly higher virulence potential than many other BV-associated species: higher adhesion capacity, higher cytotoxicity and greater ability to form a biofilm in synergistically with other BV anaerobes<sup>20,21</sup>. Furthermore, the capacity of BV-associated bacteria to form

biofilms impedes the ability of antimicrobials to treat BV and could also explain some relapses of BV observed in clinical practice<sup>18,19,22,23</sup>.

*L. crispatus* is a known biomarker of healthy vaginal microbiota<sup>24,25,26</sup>. A recent study shows that the vaginal microbiome shifted early towards an overall *Lactobacillus*-dominated profile during pregnancy<sup>27</sup>. Additionally, in the case of a high load of *Gardnerella* spp. with a low load of *A. vaginae*, *L. crispatus* under 1.2 log copies/mL was associated with the presence of BV for an additional 7% of the studied population (Figure 3).

*P. bivia* could discriminate between the absence or presence of BV among 31% of the studied population in which *A. vaginae*, *Gardnerella* spp., and *L. crispatus* were all decreased. Indeed, *P. bivia* is known to promote incident BV in synergy to *Gardnerella* spp. and *A. vaginae*<sup>28</sup>. Coinfection of mice with *Gardnerella* spp. and *P. bivia* reiterates typical characteristics of BV<sup>29</sup>. Indeed, *A. vaginae* and *Gardnerella* spp. promote vaginal proliferation of anaerobes such as *P. bivia*<sup>19,30</sup>. Muzny et al. also proposed a model of BV pathogenesis in which *A. vaginae*, *Gardnerella* spp., and *P. bivia* are sufficient to diagnose BV. In non-pregnant women, their model allowed the diagnosis of BV regardless of the diagnostic method or whether patients were symptomatic or not<sup>28</sup>. We also show that *P. bivia* could differentiate the presence or the absence of BV, especially when bacteria associated with healthy microbiota such as *L. crispatus* or bacteria associated with BV like *A. vaginae* or *Gardnerella* spp. could not. Interestingly, a recent multi-omic study showed that women who delivered preterm exhibited a decrease of *L. crispatus* and a significant increase of *A. vaginae*, *Gardnerella* spp.<sup>26</sup> indicating that the studied panel may individualize the group at risk of preterm birth.

Overall the combination of *A. vaginae*, *Gardnerella* spp., *L. crispatus*, and *P. bivia* was evaluated at 91.0%, sensitivity 77.1%, specificity 97.0% compared to the Nugent score. The low sensitivity compared to the Nugent score is the limit of the qPCR panel presented here.

However, the Nugent score is a semi-quantification based only on bacterial morphology. The scoring of morphotypes is also subjective and influenced by individual skill. Indeed, in a study including 13 experienced researchers, interobserver variability and agreement in diagnosis were observed for only 63% of cases<sup>5</sup>. Regardless, the Nugent score remains considered a gold standard used in many clinical trials for the diagnosis of BV<sup>6</sup>. Recently, a team developed a Peptide Nucleic Acid (PNA) probes by Fluorescence in situ Hybridization (FISH) methodology for BV diagnosis, which improves both sensitivity and specificity of the Nugent and provide results in 3 hours. The specificity of PNA probes for *Lactobacillus* species and *G. vaginalis* visualization associated with the calculation of the microscopic field by Nugent score allow an accurate evaluation of the bacteria present in vaginal microbiota that limit the occurrence of misleading diagnostics<sup>31</sup>. However, PNA-FISH remains semi-quantitative with unknown reproducibility due to a lack of large cohort studies.

Several studies validating molecular assays for diagnosis of BV in non-pregnant women were based on comparing qPCR to both clinical symptoms and the Nugent score<sup>9, 15, 32</sup>. One limit of our study remains not having evaluated the qPCR panel of these four bacteria in association with the symptoms of each patient. This was inherent to our cohort composed of asymptomatic pregnant women. Likewise, collection of Amsel criteria was not part of the study design. Indeed, the proportion of asymptomatic BV in pregnant women may reach 50%, limiting the sensitivity of clinical symptoms<sup>33</sup>. Additionally, the terms “symptomatic” and “asymptomatic” need a better definition, especially in pregnancy<sup>6</sup>.

In this context, the high specificity of the of four bacteria in the commercial qPCR panel screens for the absence of BV among pregnant women (high specificity). Screening for the absence of BV during pregnancy remains a crucial issue for evaluating therapeutic approach. It has been shown recently in a large cohort of pregnant women that the treatment of BV by clindamycin does not decrease the risk of preterm birth<sup>33</sup>. In this trial, treating patients was based on Nugent score results in a cohort of systematically screened parturient. Two limits

also emerge after this trial. First, the semi-quantitative Nugent score based on subjective interpretation and individual skill could have led to treat healthy women. Second, the inability to verify BV cure could have led to overestimating healthy pregnancy. Both limits could have attenuated the benefit of treatment on preterm birth rate. These two limitations are known to have hampered several studies in the field<sup>6</sup>.

Overall, the qPCR performed in the present study responds to these two limitations by implementing an easily reproducible quantitative assay that could assess either the absence of BV and or the cure of BV after treatment.

## References

1. Hill GB. The microbiology of bacterial vaginosis. *American Journal of Obstetrics and Gynecology*, 1993, 169:450–4
2. Leitich H, Bodner-Adler B, Brunbauer M, Kaider A, Egarter C, Husslein P. Bacterial vaginosis as a risk factor for preterm delivery: A meta-analysis. *American Journal of Obstetrics and Gynecology*, 2003, 189:139–47
3. Amsel R, Totten PA, Spiegel CA, Chen KCS, Eschenbach D, Holmes KK. Nonspecific vaginitis. *The American Journal of Medicine*, 1983, 74:14–22
4. Nugent RP, Krohn MA, Hillier SL. Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation. *Journal of Clinical Microbiology*, 1991, 29:297–301
5. Forsum U, Holst E, Larsson PG, Vasquez A, Jakobsson T, Mattsby-Baltzer I. Bacterial vaginosis - a microbiological and immunological enigma. Review article. *Apmis*, 2005, 113:81–90
6. Lewis AL, Laurent LC. USPSTF 2020 Recommendations on Screening for Asymptomatic Bacterial Vaginosis in Pregnancy. *JAMA*, 2020, 323:1253
7. Schwebke JR, Gaydos CA, Nyirjesy P, Paradis S, Kodosi S, Cooper CK. Diagnostic Performance of a Molecular Test versus Clinician Assessment of Vaginitis. *J Clin Microbiol*, 2018, 56:e00252-18, /jcm/56/6/e00252-18.atom
8. Cartwright CP, Lembke BD, Ramachandran K, Body BA, Nye MB, Rivers CA, Schwebke JR. Development and Validation of a Semiquantitative, Multitarget PCR Assay for Diagnosis of Bacterial Vaginosis. *Journal of Clinical Microbiology*, 2012, 50:2321–9
9. Coleman JS, Hitti J, Bukusi EA, Mwachari C, Muliro A, Nguti R, Gausman R, Jensen S, Patton D, Lockhart D, Coombs R, Cohen CR. Infectious correlates of HIV-1 shedding in the female upper and lower genital tracts: *AIDS*, 2007, 21:755–9
10. Subtil D, Brabant G, Tilloy E, Devos P, Canis F, Fruchart A, Bissinger M-C, Dugimont J-C, Nolf C, Hacot C, Gautier S, Chantrel J, Jousse M, Desseauve D, Plennevaux JL, Delaeter C, Deghilage S, Personne A, Joyez E, Guinard E, Kipnis E, Faure K, Grandbastien B, Ancel P-Y, Goffinet F, Dessein R. Early clindamycin for bacterial vaginosis in pregnancy (PREMEVA): a multicentre, double-blind, randomised controlled trial. *The Lancet*, 2018, 392:2171–9
11. Vanechoutte M, Guschin A, Van Simaey L, Gansemans Y, Van Nieuwerburgh F, Cools P. Emended description of *Gardnerella vaginalis* and description of *Gardnerella leopoldii* sp. nov., *Gardnerella piotii* sp. nov. and *Gardnerella swidsinskii* sp. nov., with delineation of 13 genomic species within the genus *Gardnerella*. *International Journal of Systematic and Evolutionary Microbiology*, 2019, 69:679–87
12. Castro J, Jefferson KK, Cerca N. Genetic Heterogeneity and Taxonomic Diversity among *Gardnerella* Species. *Trends in Microbiology*, 2020, 28:202–11
13. Jarvis JP, Rains D, Kradel SJ, Elliott J, Diamond EE, Avani-Aghajani E, Yasharpour F, Shaman JA. Diagnosing Bacterial Vaginosis with a Novel, Clinically-Actionable Molecular Diagnostic Tool. *BioRxiv* 2018. [Epub] <https://doi.org/10.1101/334177>
14. Gaydos CA, Beqaj S, Schwebke JR, Lebed J, Smith B, Davis TE, Fife KH, Nyirjesy P, Spurrell T, Furgerson D, Coleman J, Paradis S, Cooper CK. Clinical Validation of a Test for the Diagnosis of Vaginitis: *Obstetrics & Gynecology*, 2017, 130:181–9
15. Hilbert DW, Smith WL, Chadwick SG, Toner G, Mordechai E, Adelson ME, Aguin TJ, Sobel JD, Gyax SE. Development and Validation of a Highly Accurate Quantitative Real-Time PCR Assay for Diagnosis of Bacterial Vaginosis. *J Clin Microbiol*, 2016, 54:1017–24
16. Menard J, Fenollar F, Henry M, Bretelle F, Raoult D. Molecular Quantification of *Gardnerella vaginalis* and *Atopobium vaginae* Loads to Predict Bacterial Vaginosis. *CLIN*



INFECT DIS, 2008, 47:33–43

17. Hickey RJ, Forney LJ. Gardnerella vaginalis does not always cause Bacterial Vaginosis. *J Infect Dis*, 2014, 210:1682–3
18. Verstraelen H, Swidsinski A. The biofilm in bacterial vaginosis: implications for epidemiology, diagnosis and treatment. *Current Opinion in Infectious Diseases*, 2013, 26:86–9
19. Swidsinski A, Mendling W, Loening-Baucke V, Swidsinski S, Dörffel Y, Scholze J, Lochs H, Verstraelen H. An adherent Gardnerella vaginalis biofilm persists on the vaginal epithelium after standard therapy with oral metronidazole. *American Journal of Obstetrics and Gynecology*, 2008, 198:97.e1-97.e6
20. Alves P, Castro J, Sousa C, Cereija TB, Cerca N. Gardnerella vaginalis Outcompetes 29 Other Bacterial Species Isolated From Patients With Bacterial Vaginosis, Using in an In Vitro Biofilm Formation Model. *The Journal of Infectious Diseases*, 2014, 210:593–6
21. Machado A, Jefferson K, Cerca N. Interactions between Lactobacillus crispatus and Bacterial Vaginosis (BV)-Associated Bacterial Species in Initial Attachment and Biofilm Formation. *IJMS*, 2013, 14:12004–12
22. Bradshaw CS, Tabrizi SN, Fairley CK, Morton AN, Rudland E, Garland SM. The Association of Atopobium vaginae and Gardnerella vaginalis with Bacterial Vaginosis and Recurrence after Oral Metronidazole Therapy. *J INFECT DIS*, 2006, 194:828–36
23. Muli F, Struthers JK. Use of a Continuous-Culture Biofilm System To Study the Antimicrobial Susceptibilities of Gardnerella vaginalis and Lactobacillus acidophilus. *Antimicrob Agents Chemother*, 1998, 42:1428–32
24. Fredricks DN, Fiedler TL, Thomas KK, Oakley BB, Marrazzo JM. Targeted PCR for Detection of Vaginal Bacteria Associated with Bacterial Vaginosis. *Journal of Clinical Microbiology*, 2007, 45:3270–6
25. Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SSK, McCulle SL, Karlebach S, Gorle R, Russell J, Tacket CO, Brotman RM, Davis CC, Ault K, Peralta L, Forney LJ. Vaginal microbiome of reproductive-age women. *Proceedings of the National Academy of Sciences*, 2011, 108:4680–7
26. Fettweis JM, Serrano MG, Brooks JP, Edwards DJ, Girerd PH, Parikh HI, et al. The vaginal microbiome and preterm birth. *Nat Med*, 2019, 25:1012–21
27. Serrano MG, Parikh HI, Brooks JP, Edwards DJ, Arodz TJ, Edupuganti L, et al. Racioethnic diversity in the dynamics of the vaginal microbiome during pregnancy. *Nat Med*, 2019, 25:1001–11
28. Muzny CA, Blanchard E, Taylor CM, Aaron KJ, Talluri R, Griswold ME, Redden DT, Luo M, Welsh DA, Van Der Pol WJ, Lefkowitz EJ, Martin DH, Schwebke JR. Identification of Key Bacteria Involved in the Induction of Incident Bacterial Vaginosis: A Prospective Study. *The Journal of Infectious Diseases*, 2018, 218:966-978
29. Gilbert NM, Lewis WG, Li G, Sojka DK, Lubin JB, Lewis AL. Gardnerella vaginalis and Prevotella bivia Trigger Distinct and Overlapping Phenotypes in a Mouse Model of Bacterial Vaginosis. *The Journal of Infectious Diseases*, 2019, 220:1099–108
30. Hardy L, Jespers V, Abdellati S, De Baetselier I, Mwambarangwe L, Musengamana V, van de Wijgert J, Vaneechoutte M, Crucitti T. A fruitful alliance: the synergy between Atopobium vaginae and Gardnerella vaginalis in bacterial vaginosis-associated biofilm. *Sex Transm Infect*, 2016, 92:487–91
31. Machado A, Castro J, Cereija T, Almeida C, Cerca N. Diagnosis of bacterial vaginosis by a new multiplex peptide nucleic acid fluorescence in situ hybridization method. *PeerJ*, 2015, 3:e780
32. Cartwright CP, Lembke BD, Ramachandran K, Body BA, Nye MB, Rivers CA, Schwebke JR. Development and Validation of a Semiquantitative, Multitarget PCR Assay for Diagnosis of Bacterial Vaginosis. *Journal of Clinical Microbiology*, 2012, 50:2321–9
33. Klebanoff MA, Carey JC, Hauth JC, Hillier SL, Nugent RP, Thom EA, Ernest JM,

Heine RP, Wapner RJ, Trout W, Moawad A, Miodovnik M, Sibai BM, Dorsten JPV, Dombrowski MP, O'Sullivan MJ, Varner M, Langer O, McNellis D, Roberts JM, Leveno KJ. Failure of Metronidazole to Prevent Preterm Delivery among Pregnant Women with Asymptomatic *Trichomonas vaginalis* Infection. *N Engl J Med*, 2001, 345:487–93

## Figure Legends

Figure 1. Quantitative PCR result of significant bacteria relative to Nugent score  $<$  or  $\geq 7$ . P-value are represented by asterisk \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ; \*\*\*\* $P < 0.0001$ . BAVB2: *Bacterial Vaginosis Associated Bacteria-2*.

Figure 2. Quantitative PCR results of non-significant microorganisms relative to Nugent score  $<$  or  $\geq 7$ .

Figure 3. Classification and regression trees of the best combination that determines the presence or absence of bacterial vaginosis. BV-: absence of vaginosis; BV+: presence of vaginosis. Below the leaves, prevalence of respectively absence and presence of BV and percentage of the concerned population.

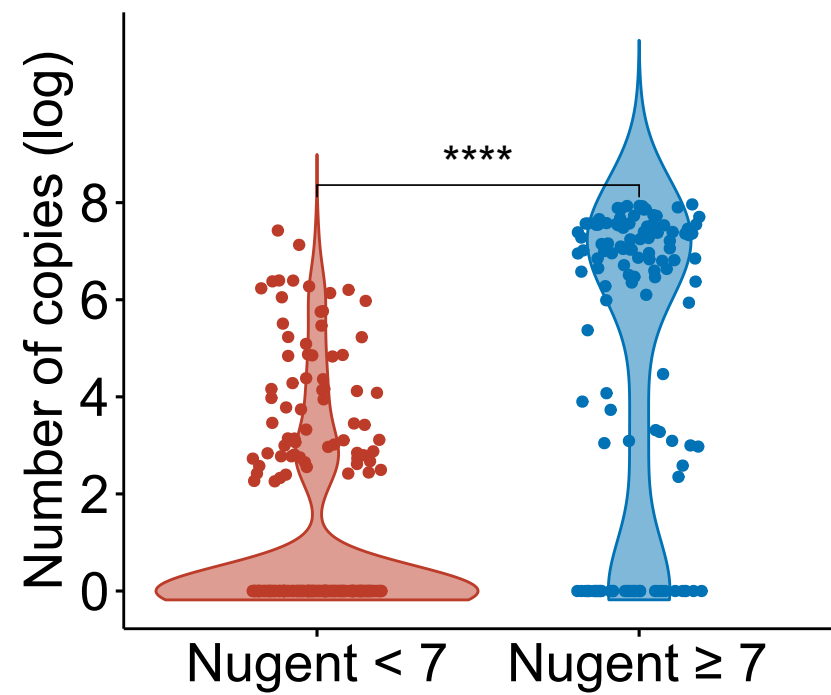
**Table 1. Qualitative PCR result of microorganisms relative to Nugent score < or >=7.**

	Nugent score < 7		Nugent score ≥7		p-value
	N=304		N=131		
	N	(%)	N	(%)	
<i>A. vaginae</i>	71	(23.4%)	90	(68.7%)	<0.001
BVAB2*	18	(5.92%)	59	(45.0%)	<0.001
<i>Gardnerella spp.</i>	112	(36.8%)	110	(84.0%)	<0.001
<i>M. curtisii</i>	5	(1.64%)	42	(32.1%)	<0.001
<i>M. mulieris</i>	0	(0.00%)	11	(8.40%)	<0.001
<i>M. hominis</i>	3	(0.99%)	41	(31.3%)	<0.001
<i>U. urealyticum</i>	118	(38.8%)	74	(56.5%)	0.001
<i>P. bivia</i>	155	(51.0%)	90	(68.7%)	0.001
<i>Megasphaera 1</i>	30	(9.87%)	74	(56.5%)	<0.001
<i>Megasphaera 2</i>	1	(0.33%)	14	(10.7%)	<0.001
<i>L. crispatus</i>	192	(63.2%)	17	(13.0%)	<0.001
<i>L. gasseri</i>	128	(42.1%)	32	(24.4%)	0.001
<i>L. iners</i>	187	(61.5%)	86	(65.6%)	0.477
<i>L. jensenii</i>	144	(47.4%)	19	(14.5%)	<0.001
<i>N. gonorrhoeae</i>	0	(0.00%)	0	(0.00%)	-
<i>C. trachomatis</i>	5	(1.64%)	5	(3.82%)	0.176
<i>T. pallidum</i>	0	(0.00%)	0	(0.00%)	-
<i>T. vaginalis</i>	1	(0.33%)	3	(2.29%)	0.084
<i>M. genitalium</i>	0	(0.00%)	1	(0.76%)	0.301
<i>H. ducrei</i>	0	(0.00%)	0	(0.00%)	-
HSV1†	0	(0.00%)	0	(0.00%)	-
HSV2†	1	(0.33%)	0	(0.00%)	1
<i>C. albicans</i>	35	(11.5%)	19	(14.5%)	0.478
<i>C. dubliniensis</i>	0	(0.00%)	0	(0.00%)	-
<i>C. glabrata</i>	3	(0.99%)	3	(2.29%)	0.372
<i>C. krusei</i>	1	(0.33%)	0	(0.00%)	1
<i>C. lusitaniae</i>	2	(0.66%)	0	(0.00%)	1
<i>C. parapsilosis</i>	1	(0.33%)	0	(0.00%)	1
<i>C. tropicalis</i>	0	(0.00%)	1	(0.76%)	0.301
<i>E. coli</i>	27	(8.88%)	21	(16.0%)	0.044
<i>E. faecalis</i>	62	(20.4%)	15	(11.5%)	0.035
<i>B. fragilis</i>	1	(0.33%)	1	(0.76%)	0.512
<i>S. aureus</i>	9	(2.96%)	2	(1.53%)	0.517
<i>S. agalactiae</i>	32	(10.5%)	25	(19.1%)	0.023

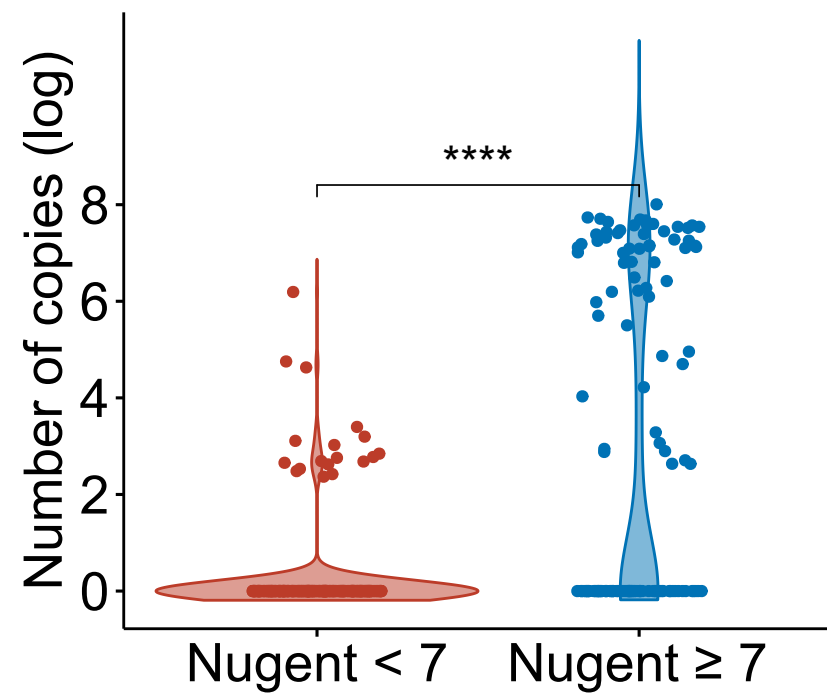
\*BAVB2: *Bacterial Vaginosis Associated Bacteria-2*

†HSV: *Herpes simplex virus*

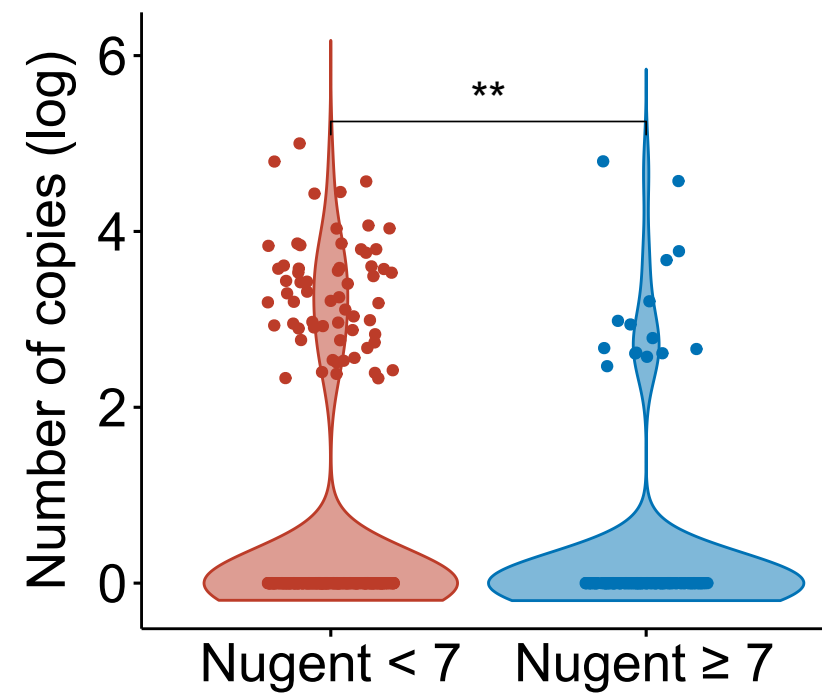
*Atopobium vaginae*



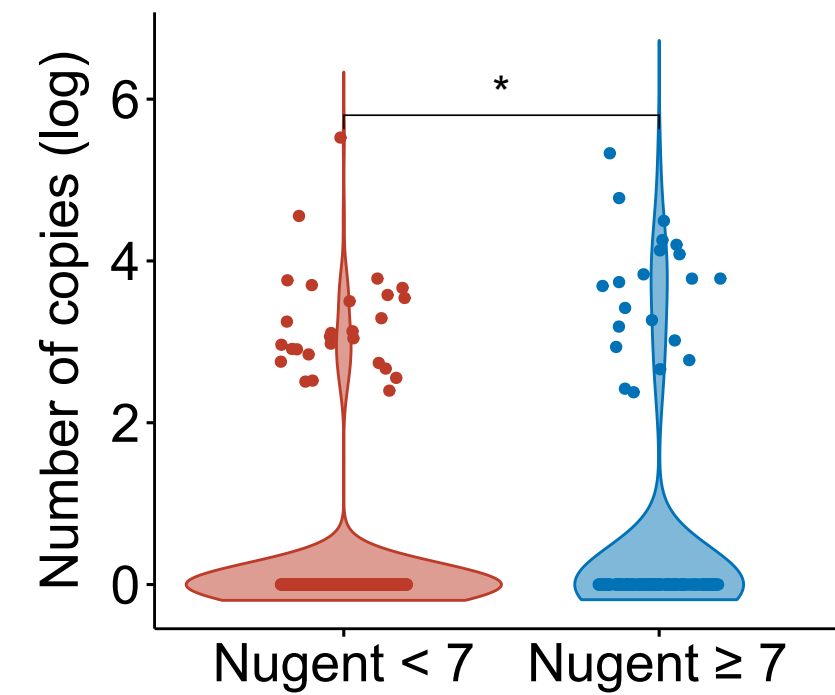
BVAB2



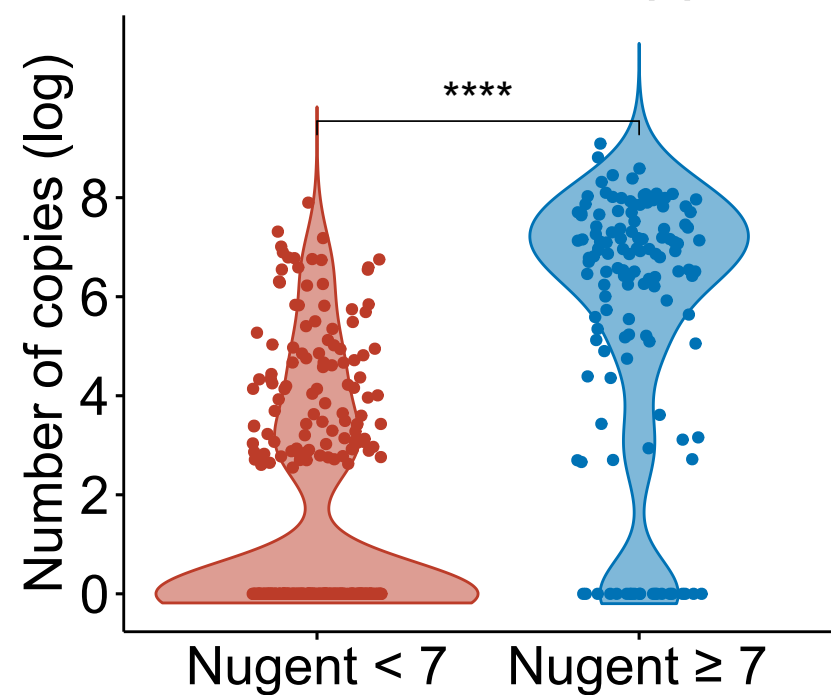
*Enterococcus faecalis*



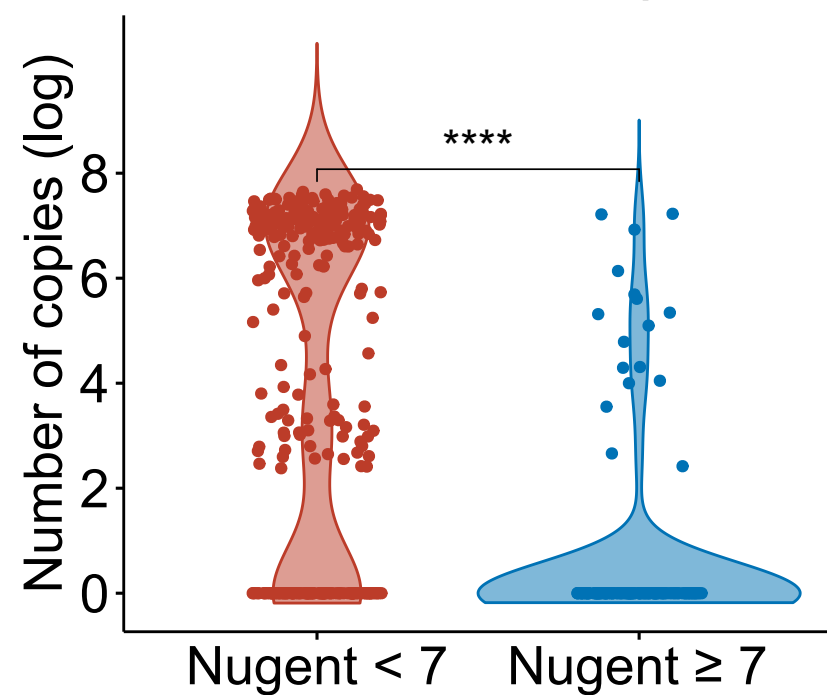
*Escherichia coli*



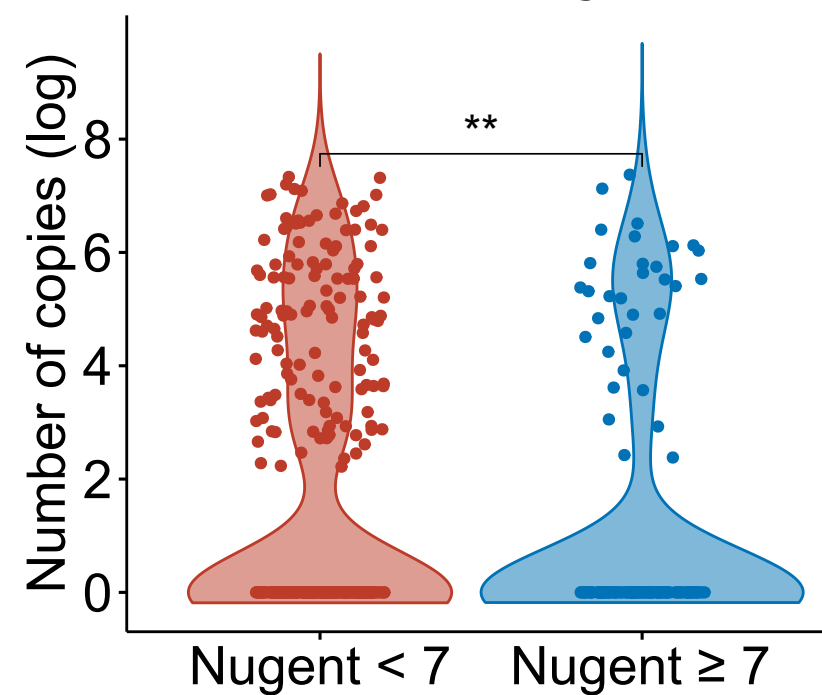
*Gardnerella* spp.



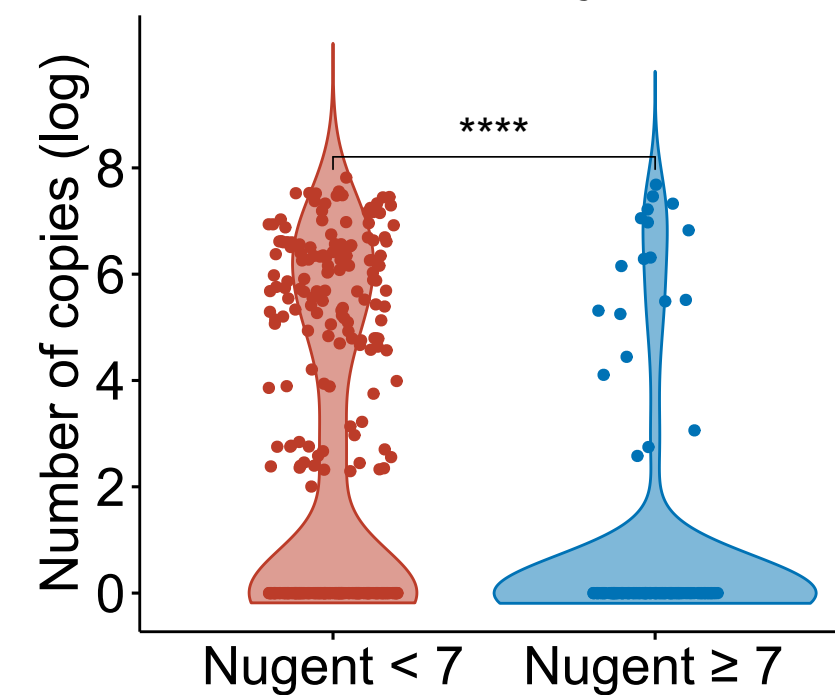
*Lactobacillus crispatus*



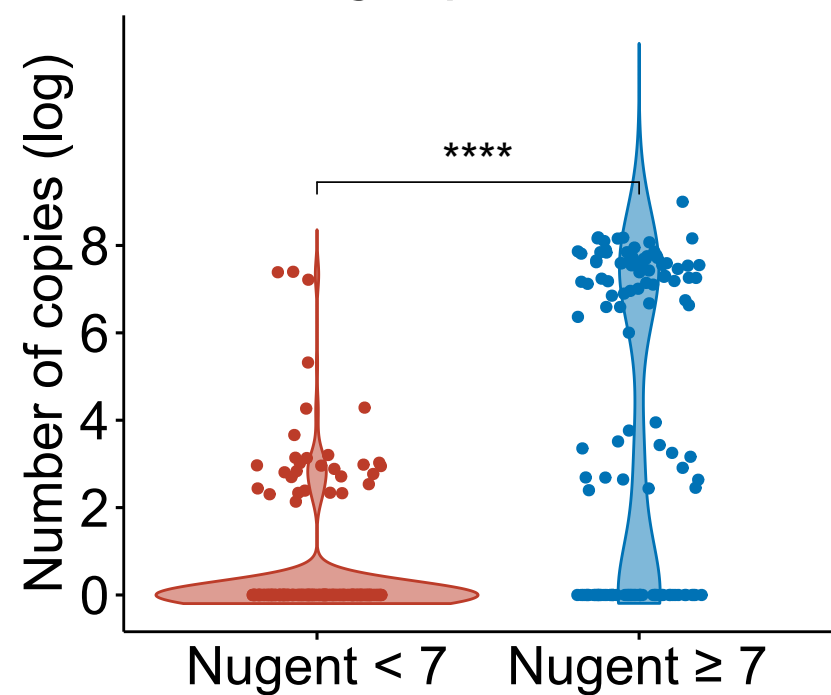
*Lactobacillus gasseri*



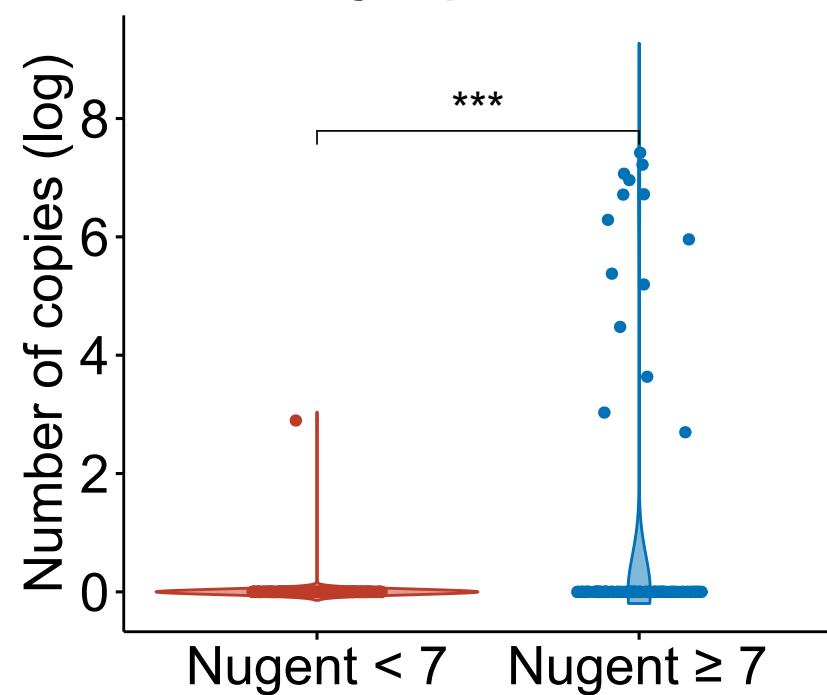
*Lactobacillus jensenii*



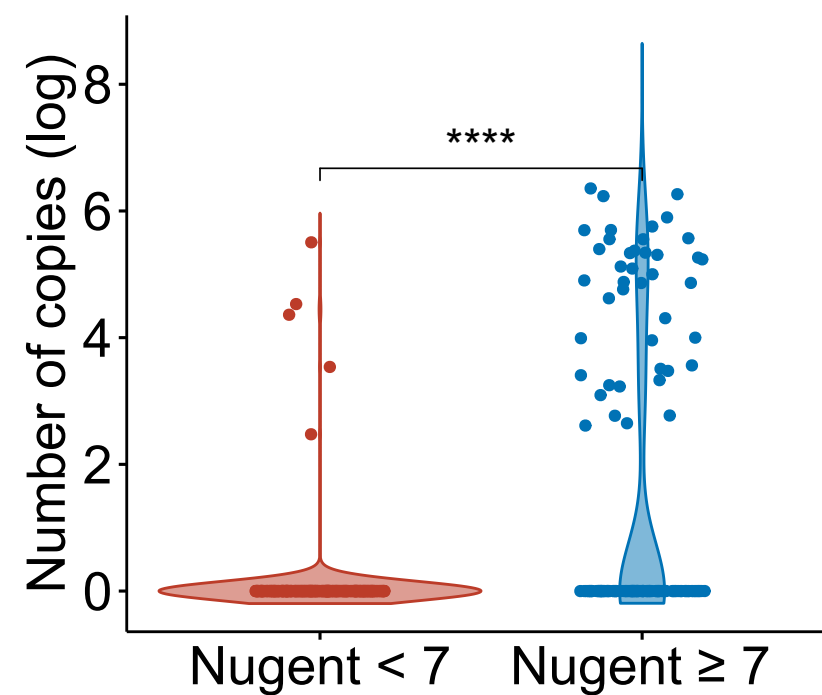
*Megasphaera* 1



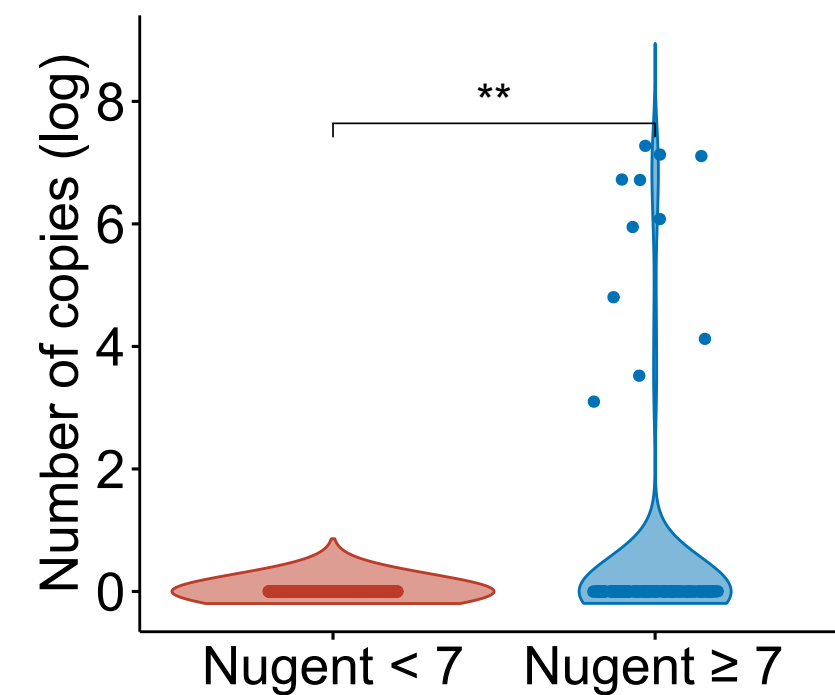
*Megasphaera* 2



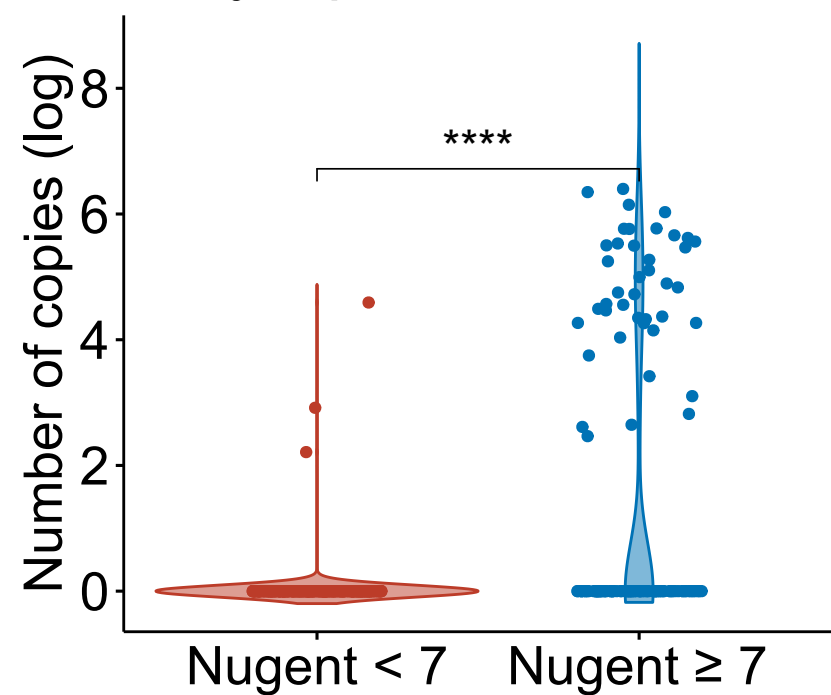
*Mobiluncus curtisii*



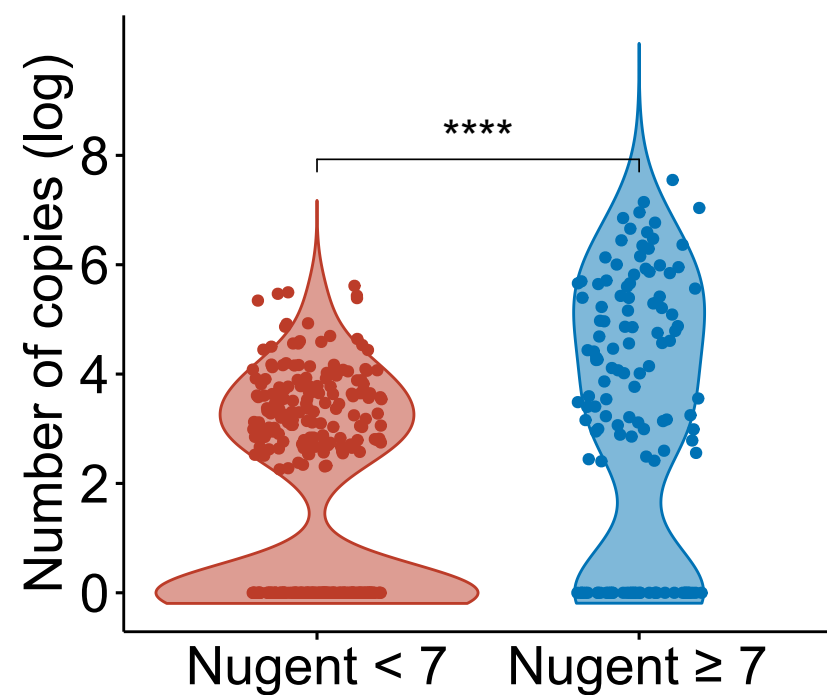
*Mobiluncus mulieris*



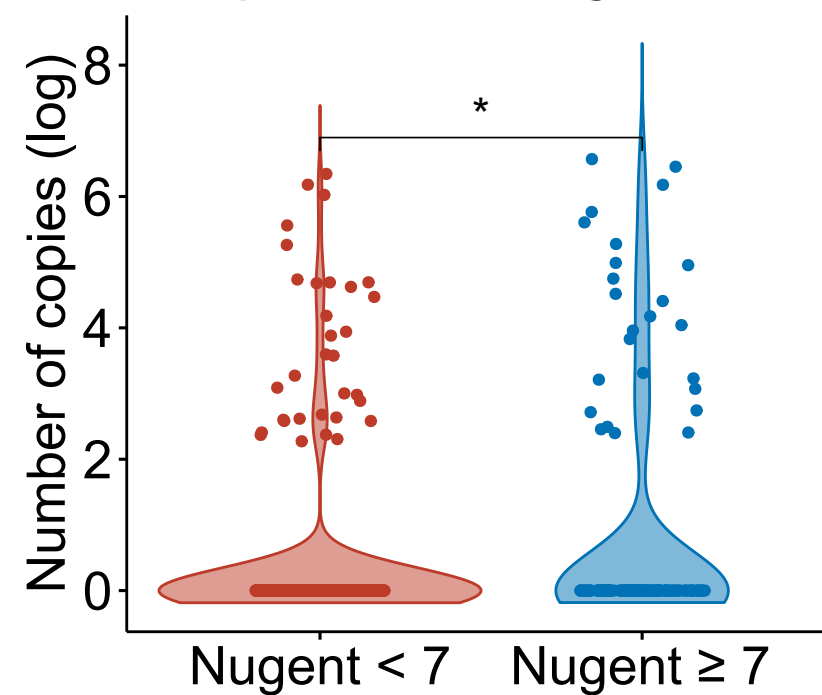
*Mycoplasma hominis*



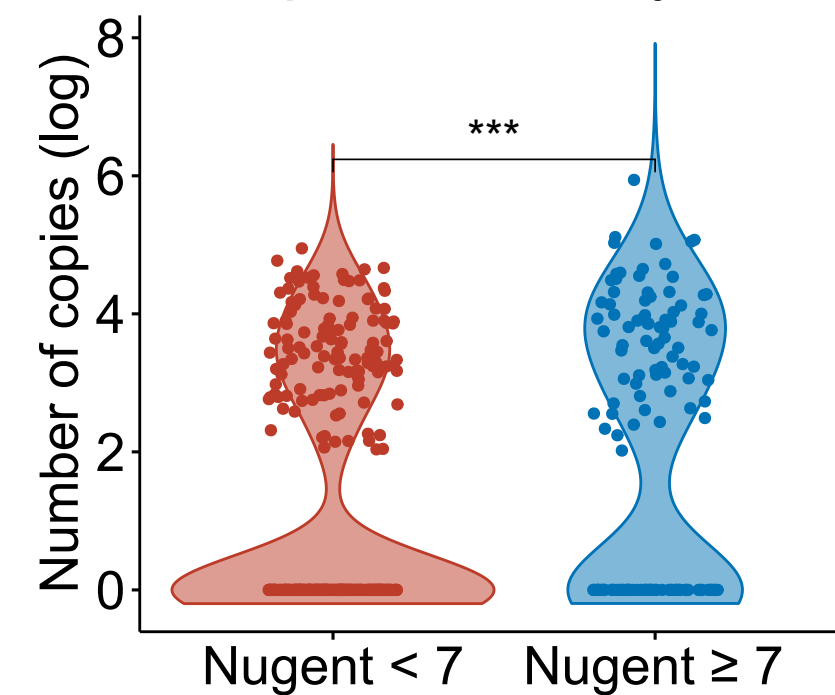
*Prevotella bivia*



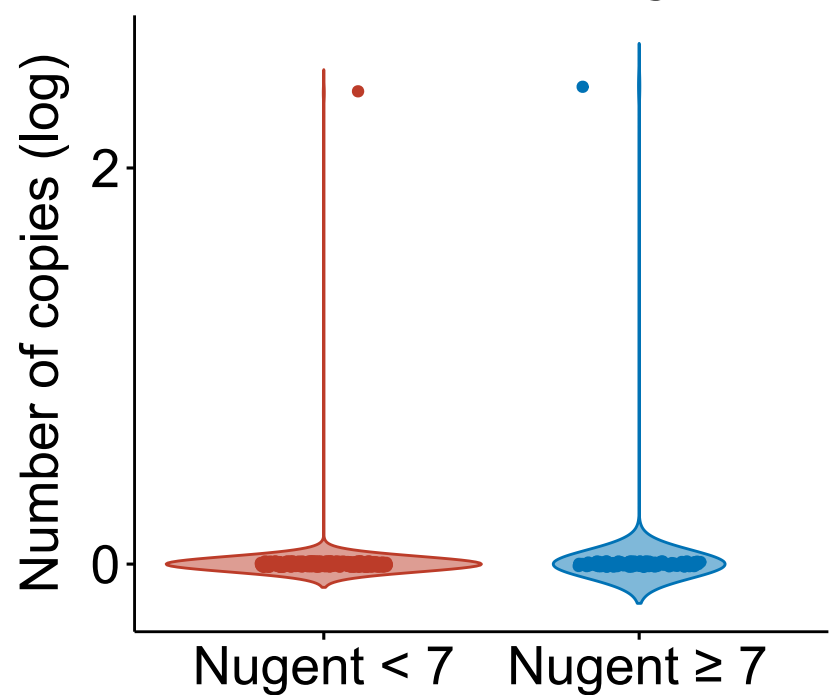
*Streptococcus agalactiae*



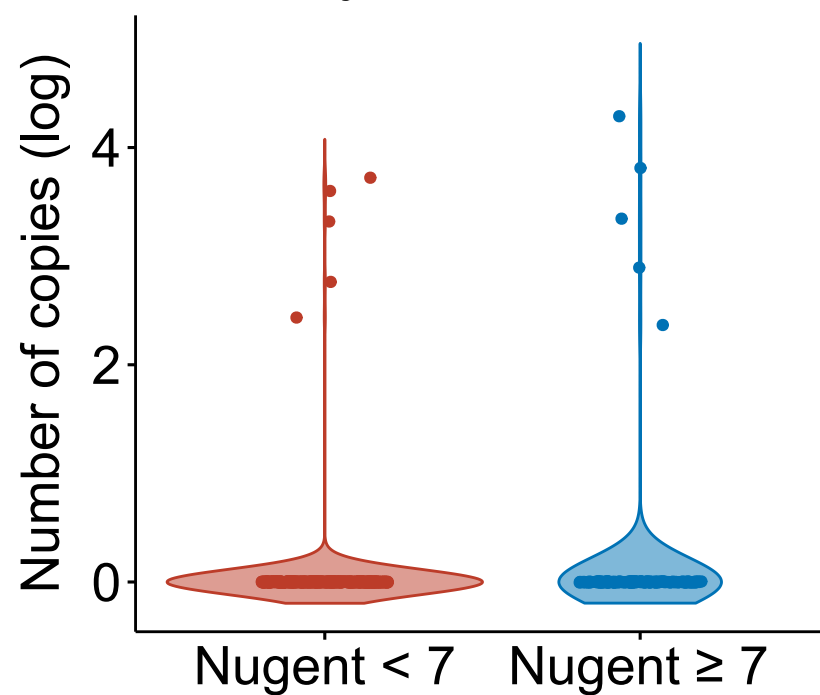
*Ureaplasma urealyticum*



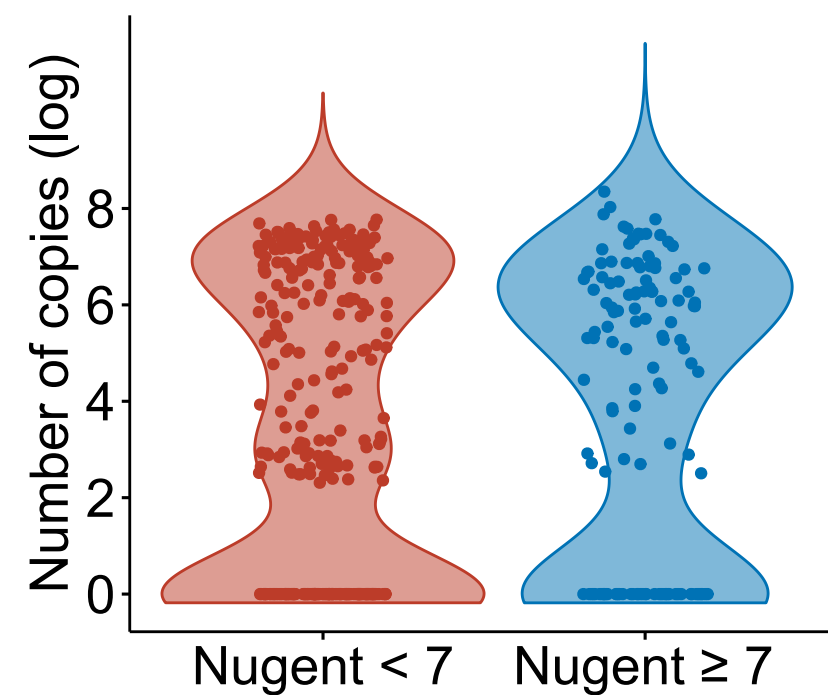
*Bacteroides fragilis*



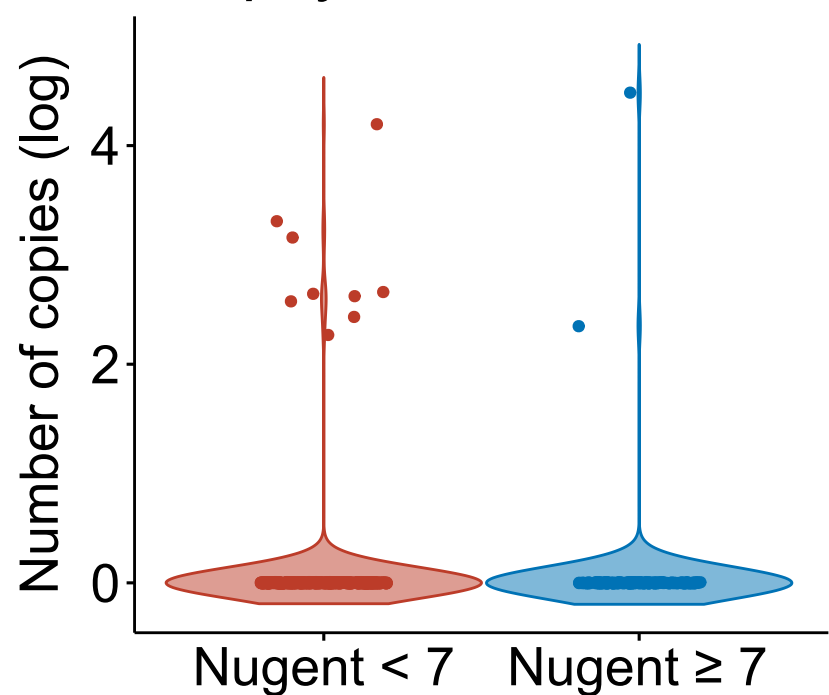
*Chlamydia trachomatis*



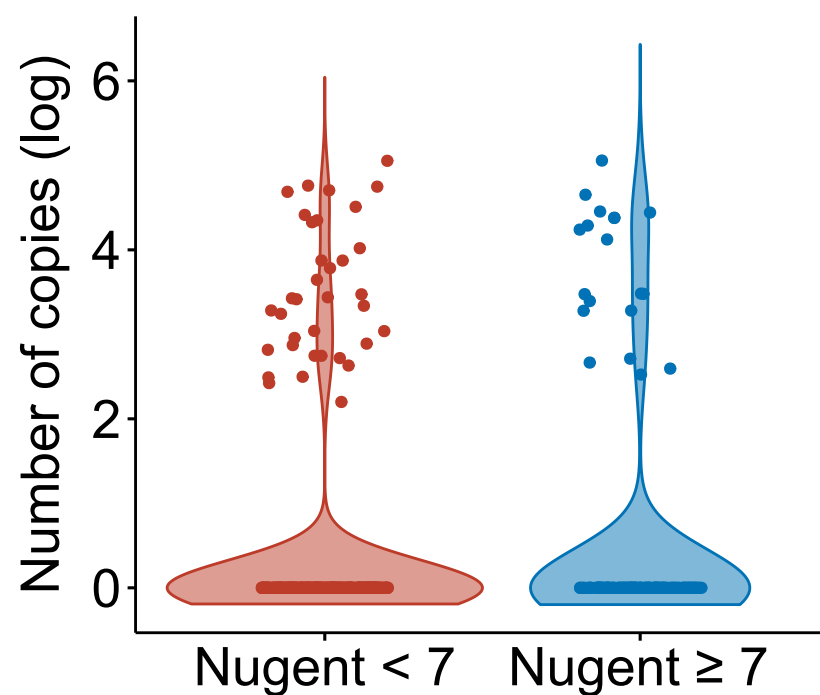
*Lactobacillus iners*



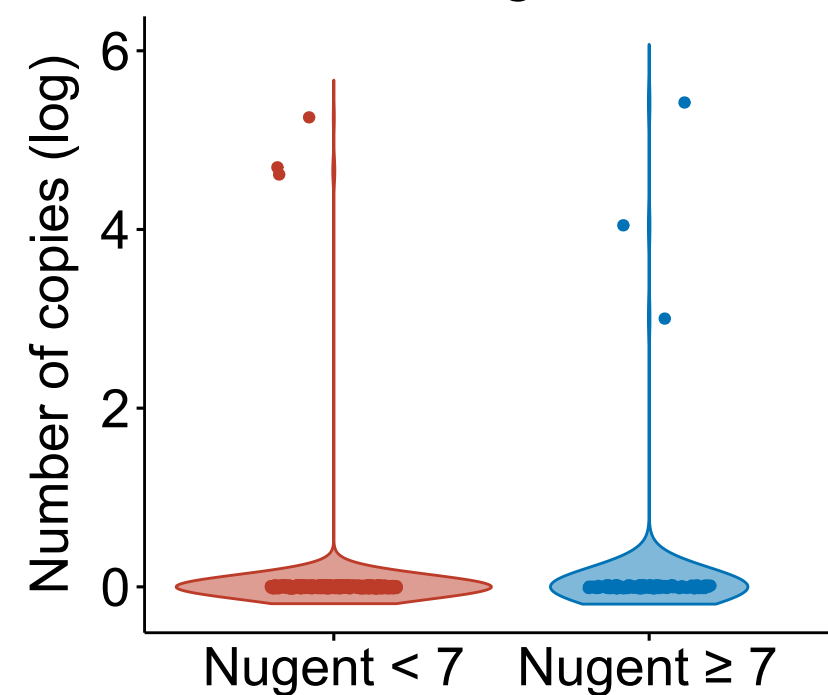
*Staphylococcus aureus*



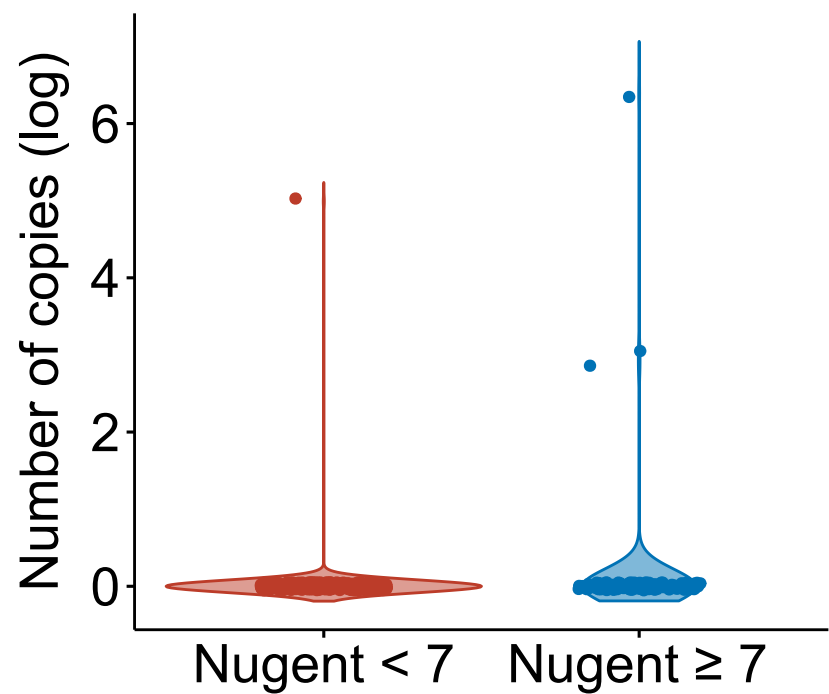
*Candida albicans*



*Candida glabrata*

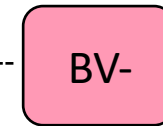


*Trichomonas vaginalis*



***Atopobium vaginae***

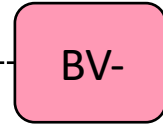
<6.4 log copies/mL      ≥6.4 log copies/mL



.70 .30  
100%

***Gardnerella spp.***

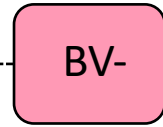
<6.4      ≥6.4



.83 .17  
83%

***Lactobacillus crispatus***

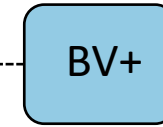
≥2.4      <2.4



.90 .10  
74%

***Lactobacillus crispatus***

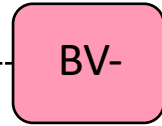
≥1.2      <1.2



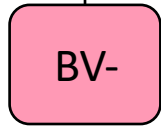
.32 .68  
9%

***Prevotella bivia***

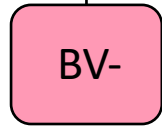
<5      ≥5



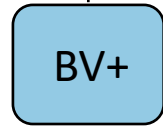
.79 .21  
31%



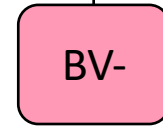
.98 .02  
43%



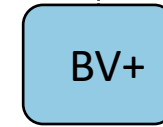
.82 .18  
30%



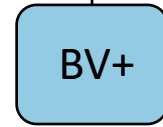
.14 .86  
2%



.70 .30  
2%



.19 .81  
7%



.03 .97  
17%